26 February 1981

MEMORANDUM FOR: Members, HTO/STIC S&T Overt Intelligence Information Processing Working Group

SUBJECT: Excerpts from Various Relevant Publications

1. The enclosed set of excerpts provides information from other publications which I believe are relevant to the working group's effort.

2. I expect that additions to this (including also revisions) will appear.

[redacted] 25X1A

25X1

[redacted]

Next 1 Page(s) In Document Exempt

Approved For Release 2004/11/04 : CIA-RDP90-00509R000100020004-1

systems and files that are now being built are designed to be shared. Unfortunately, such is not the case. Even at this moment literally tens of millions of dollars are budgeted for SAFE, a multiphased project for improving the analyst's environment that has been in the planning stage at CIA since the early 1970s.

25X1 [redacted] SAFE planning is undoubtedly impressive, but one might reasonably ask whether or not SAFE provides for resource-sharing with other members of the Intelligence Community. Unfortunately, the answer at this time must be negative. SAFE involves only CIA and DIA, and it would appear these two agencies will be limited.

25X1 [redacted] As SAFE progresses, other members of the Community seek their own solutions to the problems of information support. These local systems, which are built around minicomputers, usually include such features as text-editing, storage of reference and reporting aids, a distribution facility, provision for analysts' files, and some degree of retrospective retrieval capability - in other words, the very same capabilities that are being programmed for SAFE,... concern with Community requirements is secondary. In some instances no attempt has been made to satisfy the information needs of analysts outside one's immediate office. All over the Community the story is more or less the same.

25X1 [redacted] The Community can no longer afford a laissez faire approach to information handling. Although it is true that there will always be a requirement for special systems to meet special needs, there is a growing awareness of the commonality of the Community's interests in information distribution, storage, and retrieval.

25X1 [redacted] The basic concept of this Community system should be one of a distributed system with minimum redundancy to include the following specific features:

* All nodes of the distributed system should be linked together via communications services.

* Each of the major participants should establish and maintain its portion of the system in accordance with agreed upon standards.

* Authorized system users throughout the Community should have the capability to query any combination of the distributed data bases with a single interrogation.

* Responsibility for system development should be centralized.

* All nodes of the system should be configured identically with respect to computer manufacturer and software system employed.

* Software maintenance should be centralized.

25X1 [redacted] A certain amount of redundancy in information systems can be justified...there will always be a need for special files...to meet special needs. It is obvious, however, that we have gone far beyond any reasonable level of redundancy... It is also obvious that in fighting duplication and trying to build an effective Community-wide information-handling system, we find that the technological problems are only part, and the smaller part, of the difficulties that confront us. For us to be successful there must be a DCI-monitored commitment by the various members of the Intelligence Community to carry their part of the information load in a way that will satisfy the needs of all users.

Excerpts from Report of the Analyst Support Task Force
for the DCI's Intelligence Information Handling Committee
1 June 1979

Despite differences of production schedule, area and subject of interest, analysts expressed a common need for a rapid and easy access to all of the data available on the subjects or problems to which they were assigned.

RECOMMENDATION: Create a Community-wide, on-line, multisource bibliographic index to Intelligence Community documents, to be accessed either directly or through an information systems specialist.

Deficiencies in the present automated Community data bases-- narrow focus, complexity, incomplete data, and lack of currency-- lead analysts to duplicate data bases.

RECOMMENDATION: Press forward the ongoing effort to standardize and simplify data bases and protocols.

25X1 [redacted] The Analyst Support Task Force was convened in midSeptember 1978 by the Chairman of the DCI's Intelligence Information Handling Committee (IHC):

-- to survey the information needs of intelligence production analysts throughout the Intelligence Community,

-- to evaluate the degree to which such needs are or could be satisfied through present or improved interagency mechanisms for sharing information or data bases--both automated and manual, and

-- to recommend actions or further studies to improve the effectiveness of interagency information sharing.

25X1 [redacted] The task force focused upon answers to the following pertinent questions:

-- What information sources and data bases are now used to produce the products, and how valuable are these sources and data bases to the analyst?

-- Of those information sources and data bases identified, which are used exclusively by each analyst interviewed and which might be of value to other analysts in the Community?

-- Which data bases or information sources available to others--either within or outside the Community (e.g., academia, libraries, etc.)--would be of significant value to those analysts who do not presently have access to such data bases or sources?

-- What types of data bases that do not presently exist within the Intelligence Community

would be of future value to the analysts?

25X1 [redacted] A substantial part of each analyst's day is spent reading and evaluating a huge volume of routinely distributed source materials, largely consisting of reports that result from the collection and processing activities managed by each of the Intelligence Community agencies. ... Internal distribution procedures vary with the organization and the source; the most common, however, is to match document content--often done automaticaly--against interest profiles submitted by each potential recipient's office.

25X1 [redacted] Current intelligence--with deadlines of a few hours or at most several days--consumes 30 percent of the production time while 47 percent of the analysts' time is spent on long term or research analysis. Many analysts complained, in fact, that servicing frequent ad hoc requirements--increasingly prevents them from needed research or thorough digestion of the daily deluge of source materials.

25X1 [redacted] All source, multi-subject personal and office files are considered by the analysts interviewed to be the most valuable collection of data used--ranking above all agency, Community and non-community files presently available to the analysts.

25X1 [redacted] Many analysts believe ... that the information routinely distributed to them is adequate to do the job--or even excessive in volume. They spend so much time processing the daily source materials that they rarely have time to search out additional, hopefully, superior sources of data. As a result, those interviewed were often unaware of existing internal and external information resources, including automated files, that might be helpful to them.

Information sources which production analysts find most valuable have in common the characteristics of timeliness, convenience, and dependability. The degree of timeliness needed varies according to product deadline, but the already documented emphasis on current intelligence and quick response, ad hoc requirements enchances the need for rapid access to information. The information source must also be convenient to use--complex data retrieval systems, whether manual or automated, introduce delay and discourage analyst use.

Averaging evaluations from analysts throughout the Community, the top six sources of intelligence information ranked as follows:

1. CIA intelligence reports,

25X1 [redacted]

3. Personal contact,

4. DOD intelligence reports,

25X1 [redacted]

6. DIA and NPIC photo interpretation reports.

25X1 [redacted] Many of those interviewed were surprised at the number of the files described in the Presearch catalog of the Community's automated data

bases...They were not aware that many of these data bases exist--including some that belong to other organizations, but deal with their own subjects of expertise.

25X1 About 15 percent of the analysts interviewed complained of problems with their internal distribution system which may involve interagency shared materials....In some offices, source materials are distributed to components or supervisors and may not filter down to an interested analyst.

25X1 25X1A At CIA, the [redacted] system--an automated bibliographic index to documents--provides CIA analysts with a back-up that gives them relatively quick access to most source materials received and distributed by CIA. [redacted] 25X1A is the most widely used data base among those analysts interviewed at CIA--over 72 percent of the analysts interviewed had used it. [redacted] 25X1A is not an alternative to the distribution network, however, since documents may take up to a week to be indexed into the system, some documents do not regularly appear in the files, and the analysts must often wait several days for copies of the source documents they have identified from the automated index.

25X1 Perhaps because many of the Community's automated data bases focus on a single subject or function, relatively few analysts use any given data base--with the following exceptions. Those analysts who do use Community automated data bases, however, generally value them highly and use them intensely...

25X1 The most highly acclaimed and widely used Community data bases are those that are multi-subject or bibliographic rather than those oriented toward a specific substantive subject or area. [redacted] 25X1

25X1 [redacted] file is used regularly--at least once a month--by 36 percent of the analysts interviewed and is ranked as essential by 50 percent of those users. Another 35 percent of the [redacted] users rank the file important. 25X1

25X1 The second most widely used data base is CIA's [redacted] About 31 25X1A percent of all analysts interviewed regularly use [redacted]--the bulk of those 25X1A users at CIA and the remainder at NSA where [redacted] is accessed....in a non- 25X1A automated way.

25X1 25X1A The [redacted] systems would given analysts nearly the same capability CIA analysts presently have using the [redacted] system. The survey 25X1A findings indicate, however, that the majority of analysts would probably prefer to access such a system through an intermediary.

25X1 Over half of the analysts interviewed prefer not to access unfamiliar automated information systems directly but use them only through specialists in their reference or library components or through other analysts. These secondary users may not know which files the specialists query in order to obtain the data or documents needed, and the survey did not interview information specialists. Thus, the statistics for total automated file usage will be greater than the figures for direct automated file use by the analysts surveyed.

25X1 The interviews also reveal that more analysts would try to use automated data bases but for the deficiencies they perceive in them. These deficiencies include:

-- difficult access and/or use,

-- time consuming,

-- incomplete data,

-- lack of currency,

-- rigid and inadequate file structures.

Roughly 10 percent of the analysts interviewed complained that existing automated data bases are difficult to use, either because the system protocols are too complex and too numerous or because terminals are too few or are inconveniently located. In addition to the 39 percent who said they always preferred to access automated data bases through information systems specialists, another 20 percent of the analysts stated that they would use automated data bases themselves only if those were simple and/or were used regularly.

25X1 [redacted] ...Interviewees who use the CIRC II system of the Air Force's Foreign Technology Division (FTD) felt that the time lag of one to several weeks between making a query and receiving the hard copy document is excessive....The installation of high speed printers at agencies where CIRC II is regularly used might solve some of the document delivery tie-ups if security problems could be overcome.

Excerpts from ''Threat Evaluation Using the Open Literature''
by H.H. George and J.G. Montanaro [redacted] 25X1

25X1 [redacted] By collecting and analyzing the open literature on a US weapon system, a description can be prepared of the minimum information a dedicated adversary would be expected to know about the system. Weaknesses identified by this method are prime candidates for ECM by an adversary and could serve as a basis for electronic counter-measures improvements to the system.

Excerpts from ''TRAINER: Computer Assisted Learning and Practice for Users of DIALOG/ORBIT'' (Final Report)
by Elaine Caruso

25X1 [redacted] TRAINER is a computer assisted learning program, for scientists and other professionals, who wish to do some or all of their own literature search and reviewing, using online services. The programs, which include tutorial modules and emulations of DIALOG and ORBIT, are described: (1) as a learning environment, with descriptions of the achievement of the final test population of 15 chemist users; and (2) as computer programs designed on a DEC-10, for transportability to other computers (main frame and mini) with descriptions of sibling implementations in several other computing systems.

25X1 [redacted] A trainee will spend approximately 5 hours with the online programs to achieve a level of online searching competence sufficient to complete a search; using telecommunications networks; logging into the search service; requesting a specific file; entering search vocabulary as subject descriptor, title word, or author name; using index browsing techniques and universal character options; creating a logical statement; requesting on- or off-line printout in various formats. Strategies for optimizing search performance as appropriate to beginning users are emphasized.

25X1 [redacted] There are many arguments which can be used in support of the desirability of the option for literature users to do their own searching. These arguments can be categorized: economic, intellectual, or social. In the economics class: while the case is often made for the mediated search, as more quickly and cheaply done than the search by the literature user, this argument neglects the costs of housing and paying the intermediary, and of inconvenience and time of the user who must locate and go to the place where
25X1 service is offered. [redacted] calls this the

> ...Catch-22 for the information community. Powerful systems allow access to increasing amounts of information but only through intermediaries with whom users are reluctant to interact... Is there any sense to our present course of replacing with systems which, though automated, are more labor intensive and require an intermediary?

To continue the economic aspect of the argument we can point to the study
25X1 done by [redacted] which indicates that total use of the online system tripled when users accessed the system for themselves. Surely a tripling of the market would result in lowered costs per use. Would it also generate an intensive effort to make the systems more usable to attract a larger market that could result in even larger sales and lower costs?

Intellectual arguments can be made at several levels. Considering the actual value of completed searches at one level: users who are not present or actually running searches cannot apply their subject expertise to the selection process nor invoke their own understanding of their need in the subsequent development of the search statement. What they get is a fast batch search, which tends to be too much, too little, or not quite on target. At another level we need to consider the estrangement of the professional from the literature of his field, when we put him at a distance from the newest most effective tool for working with that literature. Part of the traditional

competence of a specialist is his knowledge of and effectiveness in using the record of research and progress in his field; and further we have evidence from a study by [redacted] that the most productive scientists and engineers prefer to do their own information gathering and that only the mediocre delegate these activities. 25X1

For social values to be served, again [redacted] has said it very well: 25X1

> Numerous observers have noted that computers appear to hold such varied and beautiful promise of expanding our alternatives and increasing freedom seem inevitably to lead to reduced options, greater rigidity, and stronger bureaucracy. Charging for services has led to outcries over the demise of the free library tradition in terms of free information, but hardly a peep is heard over free access, the freedom to search for oneself. — OR NOT TO ?

It behooves us to recognize the arguments which justify the current reality, that is, that very little searching is done by literature users; specially trained intermediaries do almost all the online searching. The fact is that most of those arguments are simply statements of what is, rather than good reasons for the continuation of the status quo; and in most instances those statements are unsupported by any evidence, or are so unqualified as to be untestable.

To summarize, the opposition to training users to do their own online searching: end users cannot or will not take time to learn how to search; they will search so infrequently that they will not be able to remember the details of the process; and searches cannot be run unless they do; the system and the databases change so much and so frequently that nothing less than dedicated study enables searchers to keep up--and they cannot search unless they do.

It was not our purpose to resolve (or even arrive at an adequate statement of) the question. We simply attempted to make it possible for alternatives to the present user-intermediary-search model to develop. We envisaged an online search environment wherein a tutorial and training component was embodied with the same accessability as the search services, and which could be used to review and refresh the searchers' skills whenever such a need arose.